

# Anomaly Detection on the Edge: Comparison of Reconstruction and Classification Based Approaches

Łukasz Grzymkowski, Tymoteusz Cejrowski

Arrow Electronics

Gdańsk, Poland

email: lukasz.grzymkowski@arrow.com,

tymoteusz.cejrowski@arrow.com

Tomasz P. Stefański

Faculty of Electronics, Telecommunications and Informatics

Gdańsk University of Technology

Gdańsk, Poland

email: tomstefa@pg.edu.pl

**Abstract**—In this work we discuss and evaluate different approaches to solving anomaly detection task when the target platform is a tiny microcontroller. We investigate modeling techniques and propose a comprehensive set of measurements to analyze performance, compute and memory requirements, and power efficiency. We run experiments to collect these measurements on platforms used in TinyML systems including Cortex-M7, Cortex-M55 and Ethos-U55 running TensorFlow Lite for Microcontrollers. The measurements are collected for an autoencoder in reconstruction-based anomaly detection and a MobileNetV2-like model trained for classification. We show which approach is more suitable depending on the system requirements and constraints. This work underscores the need for a holistic approach in selecting modeling and deployment strategies, providing empirical evidence to guide the development of efficient on-device anomaly detection systems.

**Keywords**—Edge computing, embedded software, tiny machine learning, anomaly detection, sound recognition.

## I. INTRODUCTION

Tiny Machine Learning (TinyML) enables machine learning models to be executed on resource-constrained microcontrollers and other low-power edge devices, for instance, for anomaly detection to enable on-device monitoring of heavy equipment, in automotive or industrial settings. This paper addresses the deployment of anomaly detection systems on the representative TinyML hardware. We compare distinct modeling strategies and system configurations. Through precise measurements and analysis of key metrics, this work aims to provide an insight into the current capabilities of TinyML for anomaly detection.

## II. EXPERIMENTS

We train the models for anomaly detection using the reconstruction error and the classification approaches. In the former, an autoencoder is trained on normal samples to reconstruct the input. The reconstruction error is the anomaly score. For classification, a multi-class dataset of normal samples is used to train a model, which during evaluation is used with a known class. The model negative confidence in this class is the anomaly score. We use modified baseline models, an autoencoder and a MobileNetV2-like [1], from the DCASE task 2 challenge [2]. We use MIMII [3] and ToyADMOS [4] datasets of acoustic sounds generated by industrial machines.

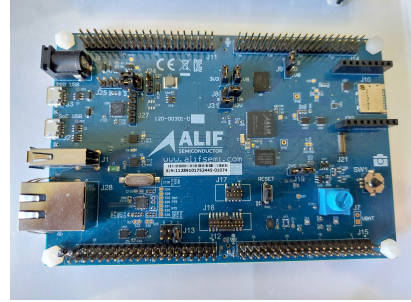


Fig. 1. DUT2 - Alif Ensemble DevKit Gen, equipped with ARM Ethos U55 Neural Processing Unit.

We evaluate the models deployed on representative devices used in TinyML systems. We use ARM Cortex-M CPUs, also test acceleration with neural processing unit, ARM Ethos-U55 (Fig. 1). We collect a comprehensive number of measurements, including inference latency and energy consumption, peak memory usage, storage for code, model parameters and MACs across different configurations.

## III. CONCLUSIONS

We conclude that to select the appropriate model and platform, a holistic approach is necessary. The decision on the model used impacts the latency, memory footprint and energy consumption, and must be made by considering technical and business requirements. In the future work, we will focus on employing neural architecture search techniques which use the proposed measurements to find optimal models or a model family for given design requirements.

## REFERENCES

- [1] M. Sandler, A. Howard *et al.*, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” in *CVPR*, 2018, pp. 4510–4520.
- [2] K. Dohi, K. Imoto *et al.*, “Description and Discussion on DCASE 2022 Challenge Task 2: Unsupervised Anomalous Sound Detection for Machine Condition Monitoring Applying Domain Generalization Techniques,” *DCASE2022*, Tech. Rep., 2022.
- [3] H. Purohit, R. Tanabe *et al.*, “MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection,” *DCASE2019*, Tech. Rep., 2019.
- [4] Y. Koizumi, S. Saito *et al.*, “ToyADMOS: A Dataset of Miniature-Machine Operating Sounds for Anomalous Sound Detection,” in *WASPAA*, 2019, pp. 308–312.